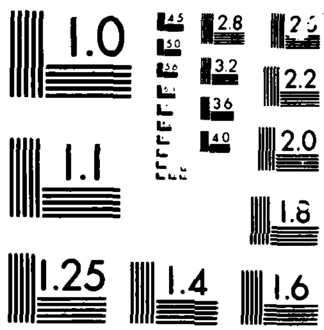
AD-A167 345 EXPONENTIAL BOUNDS OF MEAN ERROR FOR THE KERNAL ESTIMATES OF REGRESSION F..(U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS L C ZHAO DEC 85 1/1

UNCLASSIFIED TR-85-50 AFOSR-TR-86-0255 F49620-85-C-0008 F/G 12/1 NL



MICROCOPY CHART

(2)

AFOSR-TR- 86-0255

Center for Multivariate Analysis

University of Pittsburgh

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AD-A167345

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 86-0255
6a. NAME OF PERFORMING ORGANIZATION University of Pittsburgh 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code) Dept Mathmatics and Statistics Pittsburgh, PA 15260	7b. ADDRESS (City, State and ZIP Code) Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR 8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-85-C-0008
8c. ADDRESS (City, State and ZIP Code) Bolling AFB DC 20332-6448	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 61102F PROJECT NO.: 2304 TASK NO.: A5 WORK UNIT NO.:

11. TITLE (Include Security Classification)
EXPONENTIAL BOUNDS OF MEAN ERROR FOR THE KERNAL ESTIMATES OF REGRESSION FUNCTIONS

12. PERSONAL AUTHOR(S)
L.C. ZHAO

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical	FROM ______ TO ______	85 December	15

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Let (X,Y), $(X_1,Y_1),\ldots,(X_n,Y_n)$ be i.i.d. $R^r \times R$-valued random vectors with $E|Y| < \infty$, and let $Q_n(x)$ be a kernel estimate of the regression function $Q(x) = E(Y|X = x)$. In this paper, we establish an exponential bound of the mean deviation between $Q_n(x)$ and $Q(x)$ given the training sample $Z^n = (X_1,Y_1,\ldots,X_n,Y_n)$, under the conditions as weak as possible.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Major Brian W. Woodruff	22b. TELEPHONE NUMBER (Include Area Code) (202) 767-5027 22c. OFFICE SYMBOL NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

EXPONENTIAL BOUNDS OF MEAN ERROR FOR THE KERNAL ESTIMATES OF REGRESSION FUNCTIONS

L. C. ZHAO
Center for Multivariate Analysis
University of Pittsburgh

Technical Report No. 85-50

December 1985

Center for Multivariate Analysis
University of Pittsburgh
515 Thackeray Hall
Pittsburgh, PA 15260

*Research sponsored by the Air Force Office of Scientific Research (AFSC) under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation herein.

ABSTRACT

Let (X,Y), $(X_1,Y_1),\ldots,(X_n,Y_n)$ be i.i.d. $R^r \times R$-valued random vectors with $E|Y| < \infty$, and let $Q_n(x)$ be a kernel estimate of the regression function $Q(x) = E(Y|X = x)$. In this paper, we establish an exponential bound of the mean deviation between $Q_n(x)$ and $Q(x)$ given the training sample $Z^n = (X_1,Y_1,\ldots,X_n,Y_n)$, under the conditions as weak as possible.

1. INTRODUCTION

Let $(X,Y), (X_1,Y_1),\ldots,(X_n,Y_n)$ be i.i.d. $R^r \times R$-valued random vectors with $E|Y| < \infty$. To estimate $Q(x) = E(Y|X=x)$, the regression function of Y with respect to X, Nadaraya (1964) and Watson (1964) proposed a class of kernel estimates of the form

$$Q_n(x) = \sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right) Y_i \Big/ \sum_{j=1}^{n} K\left(\frac{X_j - x}{h}\right), \tag{1}$$

where K is a probability density function on R^r, and $h=h_n$ is a sequence of positive numbers. Write

$$W_{ni}(x) = K\left(\frac{X_i - x}{h}\right) \Big/ \sum_{j=1}^{n} K\left(\frac{X_j - x}{h}\right), \tag{2}$$

we define $W_{ni}(x) = 1/n$, $i=1,2,\ldots,n$, when 0/0 appears. Many scholars studied convergence problems of these estimates from various points of view. For the universal consistency, one can refer to, for example, Devroye and Wagner (1980I), Spiegelman and Sacks (1980). For the pointwise moment-consistency, see Devroye (1981). For the pointwise a.s. consistency, see Devroye (1981), Greblicki, Krzyzak and Pawlak (1984), Zhao and Fang (1985). In this paper, we study another convergence of these estimates.

Let $Z^n = (X_1,Y_1,\ldots,X_n,Y_n)$ be a training sample, $g_n = g_n(x,Z^n)$ be an estimate of $Q(x)$. In some problems, we are interested in the following mean deviation of g_n given the training sample Z^n:

$$\begin{aligned} D(g_n) &= E\{|g_n(X,Z^n) - Q(X)|\,|Z^n\} \\ &= \int_{R^r} |g_n(x,Z^n) - Q(x)|F(dx), \end{aligned} \tag{3}$$

where F denotes the distribution of X. Devroye and Wagner (1980II) proved that

$$\lim_{n\to\infty} D(Q_n) = 0 \quad \text{a.s.}$$

for the kernel estimates $Q_n(x)$ of $Q(x)$, if the following conditions are satisfied:

(i) Y is bounded,

(ii) F has a density f, (4)

(iii) K is bounded and

$$\int_{R^r} \psi(x)dx < \infty,$$

where

$$\psi(x) = \sup_{||u||>||x||} K(u), \quad x \in R^r$$

and $||.||$ is the L_2 norm or L_∞ norm on R^r,

(iv) $h_n \to 0$ and $\Sigma_n \exp(-\alpha n h_n^r) < \infty$ for any $\alpha > 0$.

In this paper, we establish an exponential bound for the above mentioned deviation of Q_n. Take $||.||$ as L_2 or L_∞ norm, and denote by $I(A)$ or I_A the indicator of set A. We establish the following

<u>Theorem</u>. Let $Q_n(x)$ b a kernel estimate defined by (1). Suppose that the following conditions are satisfied:

(i) Y is bounded.

(ii) F, the distribution of X, has a density of f. (5)

(iii) There exist positive constant α and ρ_o such that

$$K(x) \geq \alpha I(||x|| \leq \rho_o), \quad x \in R^r.$$

(iv) $h \to 0$ and $nh^r \to \infty$ as $n \to \infty$.

Then for any given $\varepsilon > 0$, we have

$$P\{D(Q_n) \geq \varepsilon\} \leq e^{-cn}. \tag{6}$$

where $C > 0$ is a constant independent of n.

Obviously, we need only to give the proof for L_∞ norm. We shall indtroduce some lemmas in section 2, and give a proof of the theorem in section 3.

2. SOME LEMMAS

For simplicity we use the following convention: $\varepsilon, \varepsilon_1, \varepsilon_2, \ldots, C, C_0, C_1, \ldots, \alpha, \beta_1, \beta_2, \delta$, ect., are all constants independent of n. $\#(A)$ denotes the cardinal of set A. $S_{x,\rho} = \{u \in R^r: ||u-x|| \leq \rho\}$. F^* and λ^* denote the outer measure generated by F and the Lebesque measure λ(on R^r), respectively. F_n denotes the empirical measure of $X^n = (X_1, \ldots, X_n)$. We now give four lemmas which are needed in the sequel.

Lemma 1(Besicovitch Covering Lemma). Let E be a bounded subset of R^r, and let K be a family of cubes covering E which contain a cube D_x with center x for each $x \in E$. Then there exist points $\{x_k\}$ in E such that

(i) $E \subset \cup D_{x_k}$,

(ii) there exists a constant σ depending only on d such that
$\sum_k I(D_{x_k}) \leq \sigma$.

For the proof, refer to Wheeden and Zygmund (1977), pp. 185-187.

Lemma 2. Let $T > 0$ be a given constant. Suppose that F has a density f. Then for any given $\varepsilon > 0$, we can choose $\beta_1 > 0$ small enough and $\beta_2 > 0$ large enough such that the set

$$E^* = \{x \in S_{o,T}: \beta_1 (2\rho)^r \leq F(S_{x,\rho}) \leq \beta_2 (2\rho)^r \quad \text{for any } \rho \in (0,1)\} \tag{7}$$

satisfies $F^*(S_{0,T} - E^*) < \varepsilon$.

Note that for any Borel-measurable set $E \subset E^*$, we have

$$\beta_1 \leq f(x) \leq \beta_2, \quad \text{for almost all } x \in E \text{(with respect to } \lambda).$$

Proof. Set

$$E_1 = \{x\epsilon S_{0,T}: \operatorname{Sup}_{0<\rho<1}\lambda(S_{x,\rho})/F(S_{x,\rho}) > 1/\beta_1\},$$

$$E_2 = \{x\epsilon S_{0,T}: \operatorname{Sup}_{0<\rho<1}F(S_{x,\rho})/\lambda(S_{x,\rho}) > \beta_2\}.$$

For any $x\epsilon E_1$ there exists a cube $S_{x,\rho}$ with $\rho\epsilon(0,1)$ such that $\lambda(S_{x,\rho}) > F(S_{x,\rho}) > F(S_{x,\rho})/\beta_1$. By Lemma 1 there exist $x_k \epsilon E_1$ and $S_{x_k,\rho_k} \triangleq S_k$ such that $\lambda(S_k) > F(S_k)/\beta_1$, $E_1 \subset \cup_k S_k$ and $\sum_k I(S_k) \leq \sigma$. Thus

$$F^*(E_1) \leq F(\cup S_k) \leq \sum F(S_k) < \beta_1 \sum \lambda(S_k)$$

$$= \beta_1 \int_{\cup S_k} \sum I(S_k) d\lambda \leq \beta_1 \sigma\lambda(\cup S_k) \leq \beta_1 \sigma \lambda(S_{o,2T}).$$

Taking $\beta_1 > 0$ small enough, we have $F^*(E_1) < \varepsilon/2$. In the same way, we have $\lambda^*(E_2) \leq F(R^d)\sigma/\beta_2 = \sigma/\beta_2$. Taking β_2 large enlugh, we can make $\lambda^*(E_2)$ small enough and, by the absolute continuity of F with respect to λ, $F^*(E_2) < \varepsilon/2$. The lemma is proved.

Fix $\delta\epsilon(0,1/2\sigma)$ and assume that $h = h_n \epsilon(0,1)$. Set

$$G_n^* = \{x\epsilon S_{0,T}: F_n(S_{x,h}) < \delta F(S_{x,h})\}. \tag{8}$$

Lemma 3. Suppose that F has a density f, $h = h_n \epsilon(0,1)$ and $nn^r \to \infty$. Then for any given $\varepsilon > 0$ we have

$$P\{F^*(G_n^*) \geq \varepsilon\} < e^{-C_1 n}.$$

Proof. By Lemma 1, there exist $x_k \epsilon G_n^*$ and $S_k = S_{x_k,h}$ such that $G_n^* \subset \cup_\kappa S_k \triangleq G$, $\sum I(S_k) \leq \sigma$. Partition R^r into sets with the form $\prod_{j=1}^{r}[(i_j-1)eh, i_j eh)$, where $i_1,\ldots,i_r = 0,\pm1,\pm2,\ldots$ and e is a fixed constant to be chosen later. Call the partition Φ, and write

$$\Phi' = \{B\epsilon\Phi: B \subset S_{0,2T}\},$$

$$\tilde{G} = \cup_{B\subset G, B\epsilon\Phi} B, \quad \tilde{S}_{0,1} = \prod_{j=1}^{r}[-1+e, 1-e),$$

$$C_{x_k} \triangleq S_k - \cup_{B\epsilon\Phi, B\subset S_k} B \subseteq x_k + h(S_{0,1} - \tilde{S}_{0,1})$$
$$\triangleq C^*_{x_k}.$$

Then

$$\lambda(C^*_{x_k}) = h^r \lambda(S_{0,1} - \tilde{S}_{0,1}) = (2h)^r[1-(1-e)^r]$$
$$\leq re\lambda(S_k).$$

Since

$$G - \tilde{G} \subset \cup_k (S_k - \cup_{B\epsilon\Phi, B\subset S_k} B) \subset \cup_k C^*_{x_k},$$

we see that

$$\lambda(G-\tilde{G}) \leq \sum_k \lambda(C^*_{x_k}) \leq re \sum_k \lambda(S_k)$$
$$\leq re \int_{\cup S_k} \sum I(S_k) d\lambda \leq re\sigma\lambda(\cup_k S_k)$$
$$\leq re\sigma\lambda(S_{0,2T}).$$

Hence we can choose e small enough to render $\lambda(G-\tilde{G})$ small enough and $F(G-\tilde{G}) < \varepsilon/4$. By (8) and the fact that $\sigma\delta < 1/2$, we get

$$F_n(\tilde{G}) \leq F_n(G) \leq \sum_k F_n(S_k) < \delta \sum_k F(S_k)$$
$$= \delta \int_{\cup S_k} \sum I(S_k) dF \leq \delta\sigma F(\cup S_k) = \delta\sigma F(G)$$
$$< \tfrac{1}{2} F(G).$$

Therefore

$$F(\tilde{G}) - F_n(\tilde{G}) > F(G) - \varepsilon/4 - \tfrac{1}{2}F(G) = \tfrac{1}{2}F(G) - \varepsilon/4,$$

and

$$F^*(G^*_n) \geq \varepsilon \quad \text{implies } F(\tilde{G}) - F_n(\tilde{G}) > \varepsilon/4.$$

For any $\mathcal{H} \subset \Phi'$, we write $\cup\mathcal{H} = \cup_{B\epsilon\mathcal{H}} B$. Then

$$\{F^*(G^*_n) \geq \varepsilon\} \subset \cup_{\mathcal{H}\subset\Phi'} \{F(\cup\mathcal{H}) - F_n(\cup\mathcal{H}) > \varepsilon/4\}.$$

Assume that $\varepsilon \epsilon (0,1)$. By Hoeffding's inequality,

$$\mathrm{Sup}_A P\{F(A)-F_n(A)>\varepsilon/4\} \le \mathrm{Sup}_A \exp\{-n(\varepsilon/4)^2/[2F(A)+\varepsilon/4]\}$$

$$\le \exp\{-n\varepsilon^2/[16(2+\varepsilon/4]\} \le \exp(-n\varepsilon^2/36).$$

Noticing that $\#\{H: H\subset\Phi'\} \le 2^{C_0 h^{-r}}$ and $h^{-r} = o(n)$, we get

$$P\{F^*(G_n^*)\ge\varepsilon\} \le 2^{C_0 h^{-r}} \mathrm{Sup}_A P\{F(A)-F_n(A)\ge\varepsilon/4\}$$

$$\le \exp(-C_1 n),$$

and the lemma is proved.

Lemma 4. Suppose that $\int_{R^r}|g(x)|^p F(dx)<\infty$ for some $p > 0$, then

$$\lim_{h\to 0}\int_{S_{x,h}}|g(u)-g(x)|^p F(du)/F(S_{x,h}) = 0$$

for almost all x (with respect to F).

Refer to Wheeden and Zygmund (1977), p. 191, example 20.

3. PROOF OF THE THEOREM

Assume that $|Y| \leq M$. Without loss of generality, we can assume that $\rho_o = 1$ in (5)(iii). It is enough to prove that for each fixed $T > 0$,

$$P\{\int_{S_{0,T}} |Q_n(x) - Q(x)| F(dx) > \varepsilon\} < e^{-cn}. \tag{9}$$

By Lemma 2, there exist $\beta_1 = \beta_1(\varepsilon)$, $\beta_2 = \beta_2(\varepsilon)$ and a compact set $E \subset E^*$ such that $F(S_{o,T}-E) < \varepsilon/8M$, wher E^* is defined by (7). Hence

$$\int_{S_{0,T}-E} |Q_n(x) - Q(x)| F(dx) \leq 2MF(S_{0,T}-E) < \varepsilon/4.$$

Fix $\delta \in (0,1/2\sigma)$. By Lemma 3, there exists a compact set H_n such that

$$H_n \subset \{x \in S_{0,T}: F_n(S_{x,h}) \geq \delta F(S_{x,h})\}, \tag{10}$$

and

$$P\{F(S_{0,T}-H_n) \geq \varepsilon/(8M)\} < e^{-C_1 n} \tag{11}$$

Hence

$$P\{\int_{S_{0,T}-H_n} |Q_n(x) - Q(x)| F(dx) > \varepsilon/4\} < e^{-C_1 n}.$$

Therefore, we need only to prove that

$$P\{\int_{H_n \cap E} |Q_n(x) - Q(x)| F(dx) > \varepsilon/2\} < e^{-C_2 n}. \tag{12}$$

For $x \in H_n \cap E$, by (5)(iii), (7) and (10), we have

$$\sum_{j=1}^{n} K(\frac{X_j - x}{h}) \geq n\alpha F_n(S_{x,h}) \geq n\alpha\delta F(S_{x,h})$$

$$\geq n\alpha\delta\beta_1 2^r h^r,$$

and $f(x) \leq \beta_2$. Write $C_3 = \beta_2/(\alpha\delta\beta_1 2^r)$, we see that

$$\int_{H_n \cap E} |Q_n(x) - Q(x)| f(x) dx$$

$$\leq C_3 (nh^r)^{-1} \int_{H_n \cap E} |\sum_{i=1}^{n} K(\frac{X_i - x}{h})(Y_i - Q(x))| dx.$$

There exist finite positive constants $m, a_1, \ldots, a_m$ and disjoint regular cubes $A_1, \ldots, A_m$ such that $K^*(x) = \sum_{i=1}^{m} a_i I_{A_i}(x)$ satisfies

$$\int_{R^r} |K(x) - K^*(x)|dx < \varepsilon/(8C_3M).$$

Here a regular cube means a r-fold product of one-dimensional compact intervals.

Thus

$$P\{(nh^r)^{-1}\int_{H_n \cap E} |\sum_{i=1}^{n} (K(\frac{X_i-x}{h}) - K^*(\frac{X_i-x}{h}))(Y_i-Q(x))|dx$$

$$\leq 2C_3M(nh^r)^{-1} \sum_{i=1}^{n} \int |K(\frac{X_i-x}{h}) - K^*(\frac{X_i-x}{h})|dx$$

$$\leq 2C_3M\int |K(x) - K^*(x)|dx < \varepsilon/4.$$

Take $\varepsilon_1 = \varepsilon/(4C_3)$. To prove (12), it is enough to prove that

$$P\{(nh^r)^{-1}\int_{H_n \cap E} |\sum_{i=1}^{n} K(\frac{X_i-x}{h})(Y_i-Q(x))|dx > \varepsilon_1\}$$
$$< e^{-C_2 n}.$$

It is sufficient for any $\varepsilon_2 > 0$ and any regular cube A to prove that

$$P\{(nh^r)^{-1}\int_{H_n \cap E} |\sum_{i=1}^{n} I_{x+hA}(X_i)(Y_i - Q(X_i))|dx \geq \varepsilon_2\} \tag{13}$$
$$< e^{-C_3 n},$$

and

$$P\{(nh^r)^{-1}\int_{H_n \cap E} |\sum_{i=1}^{n} I_{x+hA}(X_i)(Q(X_i)-Q(x)|dx \geq \varepsilon_2\}$$
$$< e^{-C_3 n}. \tag{14}$$

We proceed to prove (13). To this end, we construct the pratition Φ of R^r mentioned in the proof of Lemma 3. Assume that $A = \prod_{i=1}^{r} [x_i, x_i+a_i]$ and $\min a_i \geq 2e$. Set $\tilde{A} = \prod_{i=1}^{r} [x_i+e, x_i+a_i-e)$. $A_x = \cup_{B \in \Phi, B \subseteq x+hA} B$,

$$C_x = x + hA - A_x \subset x + h(A-\tilde{A}) = C_x^*.$$

It is easy to see that we can make $\lambda(A-\tilde{A})$ arbitrarily small by choosing e small enough. We have

$$\begin{aligned}
&(nh^r)^{-1}\int_{H_n\cap E} \Big|\sum_{i=1}^n I_{x+hA}(X_i)(Y_i-Q(X_i)\Big|dx \\
&\le (nh^r)^{-1}\int_{H_n\cap E} \Big|\sum_{i=1}^n I_{A_x}(X_i)(Y_i-Q(X_i))\Big|dx \\
&\quad + 2Mh^{-r}\int F_n(C_x^*)dx \\
&\le (nh^r)^{-1}\int_{H_n\cap E}\sum_{B\epsilon\Phi, B\subseteq x+hA} \Big|\sum_{i=1}^n I_B(X_i)(Y_i-Q(X_i))\Big|dx \\
&\quad + 2M\lambda(A-\tilde{A}).
\end{aligned} \tag{15}$$

Here we use the fact that $\int\nu(x+hD)dx = h^r\lambda(D)$ for any r-dimensional probability measure ν and any Borel set $D\subset R^r$. We can choose e such that $2M\lambda(A-\tilde{A})<\varepsilon_2/2$. Note that for $B\epsilon\Phi$, $\lambda\{x.B\subseteq x+hA\} \le C_4h^r$, and $\cup_{x\epsilon H_n\cap E}\{x+hA\}\subset S_{0,2T}$ for small h. Hence, for large n, we have

$$(15) \le C_4n^{-1}\sum_{B\epsilon\Phi}\Big|\sum_{i=1}^n I_B(X_i)(Y_i-Q(X_1))\Big| + \varepsilon_2/2.$$

Set $\varepsilon_3 = \varepsilon_2/(4C_4)$. To prove (13), we need only to prove that

$$P\{\sum_{B\epsilon\Phi}\Big|\sum_{i=1}^n I_B(X_i)Y_i - n\int_B Q(x)dF\Big|\ge n\varepsilon_3\} < e^{-C_5n}, \tag{16}$$

$$P\{\sum_{B\epsilon\Phi}\Big|\sum_{i=1}^n I_B(X_i)Q(X_i) - n\int_B Q(x)dF\Big|\ge n\varepsilon_3\} < e^{-C_5n}. \tag{17}$$

Let N be a Poisson random variable with mean value n, which is independent of $(X_1,Y_1),(X_2,Y_2),\dots$. In the sequel, we use $\sum'$ for $\sum_{B\epsilon\Phi}$. Notice

That B's are disjoint, we see that for $t_B \epsilon (-\infty,\infty)$,

$$E\{\exp \sum{}' t_B \sum_{i=1}^{N} I_B(X_i)Y_i)\}$$

$$= \sum_{\ell=0}^{\infty} \frac{n^{\ell} e^{-n}}{\ell!} [E\{\exp(\sum{}' t_B I_B(X_1)Y_1)\}]^{\ell}$$

$$= \exp\{n \sum{}' E[I_B(X_1)(e^{t_B Y_1} - 1)]\},$$

So that, $\{|\sum_{i=1}^{N} I_B(X_i)Y_i - n\int_B Q(x)dF|, B\epsilon\Phi'\}$ is a group of mutually independent variables. Set

$$Z(B,N) = \sum_{i=1}^{N} I_B(X_i)Y_i - n\int_B Q(x)dF.$$

for $t > 0$, notice that $e^t - t \geq e^{-t} + t$, we have

$$P\{\sum{}' |Z(B,N)| \geq \tfrac{1}{2}n\varepsilon_3\} \leq \exp(-\tfrac{1}{2}tn\varepsilon_3)E\exp(t\sum{}' |Z(B,N)|)$$

$$= \exp(-\tfrac{1}{2}tn\varepsilon_3)\Pi_{B\epsilon\Phi'} E\{\exp(t|Z(B,N)|)\}$$

$$\leq \exp(-\tfrac{1}{2}tn\varepsilon_3)\Pi_{B\epsilon\Phi'} [E\exp(tZ(B,N) + E\exp(-tZ(B,N))]$$

$$= \exp(-\tfrac{1}{2}tn\varepsilon_3)\Pi_{B\epsilon\Phi'} \{\exp[nEI_B(X_1)(e^{tY_1} - tY_1 - 1)] \tag{18}$$

$$+ \exp[nEI_B(X_1)(e^{-tY_1} + tY_1 - 1)]\}$$

$$\leq \exp(-\tfrac{1}{2}tn\varepsilon_3)\Pi_{B\epsilon\Phi'} \{2\exp[nEI_B(X_1)(e^{t|Y_1|} - t|Y_1| - 1]\}$$

$$\leq \exp(-\tfrac{1}{2}tn\varepsilon_3)2^{C_0 h^{-r}} \exp\{n\sum{}' E[I_B(X_1)(e^{tM} - tM - 1)]\}$$

$$\leq \exp(-\tfrac{1}{2}tn\varepsilon_3)e^{o(n)}\exp\{n(e^{tM} - tM - 1)\}.$$

Take $t\epsilon(0,1/M)$, we see that $e^{tM} - tM - 1 \leq t^2M^2$. Hence, we can take $t > 0$ such that

$$P\{\sum{}' |\sum_{i=1}^{N} I_B(X_i)Y_i - n\int_B Q(x)dF| \geq \tfrac{1}{2}n\varepsilon_3\} < e^{-C_6 n}. \tag{19}$$

Write $\Delta = (X_1, X_2, \ldots)$. By Jensen's inequality, for $t > 0$ we have

$$P\{\sum{}' \left| \sum_{i=1}^{N} I_B(X_i)Q(X_i) - n\int_B Q(x)dF \right| \geq \tfrac{1}{2}n\varepsilon_3\}$$

$$= P\{\sum{}' |E(Z(B,N)|\Delta)| \geq \tfrac{1}{2}n\varepsilon_3\}$$

$$\leq P\{E(\sum{}' |Z(B,N)||\Delta) \geq \tfrac{1}{2}n\varepsilon_3\}$$

$$\leq \exp(-\tfrac{1}{2}tn\varepsilon_3)E\{\exp(tE[\sum{}' |Z(B,N)||\Delta])\}$$

$$\leq \exp(-\tfrac{1}{2}tn\varepsilon_3)E\{E[\exp(t\sum{}' |Z(B,N)|)|\Delta]\}$$

$$= \exp(-\tfrac{1}{2}tn\varepsilon_3)E[\exp(t\sum{}' |Z(B,N)|)].$$

By (18) and (19), we can take $t > 0$ such that

$$P\{\sum{}' \left| \sum_{i=1}^{N} I_B(X_i)Q(X_i) - n\int_B Q(x)dF \right| \geq \tfrac{1}{2}n\varepsilon_3\} < e^{-C_6 n}. \tag{20}$$

Note that

$$\sum{}' \left| \sum_{i=1}^{N} I_B(X_i)Y_i - \sum_{i=1}^{n} I_B(X_i)Y_i \right| \leq M|N-n|,$$

$$\sum{}' \left| \sum_{i=1}^{N} I_B(X_i)Q(X_i) - \sum_{i=1}^{n} I_B(X_i)Q(X_i) \right| \leq M|N-n|,$$

we have

$$P\{\sum{}' \left| \sum_{i=1}^{N} I_B(X_i)Y_i - \sum_{i=1}^{n} I_B(X_i)Y_i \right| \geq \tfrac{1}{2}n\varepsilon_3\} \tag{21}$$

$$\leq P\{|N-n| \geq n\varepsilon_3/(2M)\} < e^{-C_7 n},$$

$$P\{\sum{}' \left| \sum_{i=1}^{N} I_B(X_i)Q(X_i) - \sum_{i=1}^{n} I_B(X_i) \right| \geq \tfrac{1}{2}n\varepsilon_3\}$$

$$< e^{-C_7 n}. \tag{22}$$

From (19) - (22), (16) and (17) follows, and (13) is proved. It remains to prove (14). To this end, we need only to prove

$$P\{\sum_{i=1}^{n} Z(X_i) \geq n\varepsilon_2\} < e^{-C_3 n}, \tag{23}$$

where

$$\begin{aligned} Z(u) &= h^{-r}\int_{R^r} I_{x+hA}(u)|Q(u) - Q(x)|dx \quad . \\ &\leq 2Mh^{-r}\int I_A(\frac{u-x}{h})dx = 2M\lambda(A) \triangleq C_8 \end{aligned} \tag{24}$$

Hence, we have

$$|\sum_{i=1}^{N} Z(X_i) - \sum_{i=1}^{n} Z(X_i)| \leq C_8|N-n|,$$

and

$$\begin{aligned} &P\{\sum_{i=1}^{N} Z(X_i) - \sum_{i=1}^{n} Z(X_i)| \geq \tfrac{1}{2}n\varepsilon_2\} \\ &\leq P\{|N-n| \geq n\varepsilon_2/(2C_8)\} < e^{-C_9 n}. \end{aligned} \tag{25}$$

For $t > 0$, we have

$$\begin{aligned} &P\{\sum_{i=1}^{N} Z(X_i) > \tfrac{1}{2}n\varepsilon_2\} \leq \exp(-\tfrac{1}{2}tn\varepsilon_2)E\{\exp(t\sum_{i=1}^{N} Z(X_i)\} \\ &= \exp\{-\tfrac{1}{2}nt\varepsilon_2 + n\int(e^{tZ(u)}-1)F(du)\}. \end{aligned} \tag{26}$$

Take $t \in (0,1/C_8)$. By $0 \leq tZ(u) \leq 1$ we get

$$n\int(e^{tZ(u)}-1)F(du) \leq 2nt\int Z(u)F(du) \tag{27}$$

Take $\rho > 0$ so large that $A \subset S_{0,\rho}$. Then, by Lemma 4, we have

$$\begin{aligned} Z(u) &= h^{-r}\int_{u-hA}|m(x) - m(u)|dx \\ &\leq \lambda(S_{0,\rho})\int_{S_{u,h\rho}}|m(x) - m(u)|dx/\lambda(S_{u,h\rho}) \end{aligned}$$

$\to 0$ as $h \to 0$, for almost all $x(\lambda)$.

In view of (24), from the dominated convergence theorem, we see that

$$\lim_{h\to 0}\int Z(u)F(du) = 0. \tag{28}$$

By (26) - (28), we can take $t > 0$ sufficiently small such that

$$P\{\sum_{i=1}^{N} Z(X_i) > \tfrac{1}{2}n\varepsilon_2\} \leq \exp(-\tfrac{1}{2}nt\varepsilon_2 + o(nt))$$

$$< e^{-C_{10}n}.$$

From (25) and (29), we obtain (23), and (14) follows. Up to now, the theorem is proved.

REFERNCES

[1] Devroye, L.P. (1981). On the Almost Everywhere Convergence of Nonparametric Regression Function Estimates. Ann. Statist., 9, 1310-1319.

[2] Devroye, L.P. and Wagner, T.J. (1980I). Distribution Free Consistency Results in Nonparametric Discriminatin and Regression Function Estimation. Ann. Statist., 8, 231-239.

[3] Devroye, L.P. and Wagner, T.J. (1980II). On the L_1 Convergence of Kernel Estimators of Regression Functions With Applications in Discriminatin. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 51, 15-25.

[4] Greblicki, W., Krzyak, A. and Pawlak, M. (1984). Distribtuion-free Pointwise Consistency of Kernel Regression Estimate. Ann. Statist., 12, 1570-1575.

[5] Nadaraya, E.A. (1964). On Estimating Regression. Theor. Probability and Appl., 9, 141-142.

[6] Spiegelman, C. and Sacks, J. (1980). Consistent Window Estimation in Nonparametric Regression. Ann. Statist., 8, 240-246.

[7] Watson, G.S. (1964). Smooth Regression Analysis. Sankhya, Ser. A, 26, 359-372.

[8] Wheeden, R.L. and Zygmund, A. Measure and Integral, Marcel Dekker, New York, 1977.

[9] Zhao, Lincheng and Fang, Zhaoben (1985). Strong Convergence of Kernel Estimates of Nonparametric Regression Functions. Chin. Ann. of Math., Ser. B, 6, 147-155.

END
FILMED

6-86

DTIC